

The Maximum Number of Parameters for the Hausman Test When the Estimators are from Different Sets of Equations*

Kazumitsu Nawata

Graduate School of Engineering
University of Tokyo
Japan

Michael McAleer

Department of Quantitative Finance
National Tsing Hua University
Taiwan
and
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Department of Quantitative Economics
Complutense University of Madrid

December 2013

* This paper was supported by a Grant-in-Aid for Scientific Research “Analyses of the Large Scale Medical Survey Data and the Policy Evaluations in Japan (Grant Number: 24330067)” of the Japan Society of Science for the first author, and Australian Research Council and the National Science Council, Taiwan for the second author.

Abstract

Hausman (1978) developed a widely-used model specification test that has passed the test of time. The test is based on two estimators, one being consistent under the null hypothesis but inconsistent under the alternative, and the other being consistent under both the null and alternative hypotheses. In this paper, we show that the asymptotic variance of the difference of the two estimators can be a singular matrix. Moreover, in calculating the Hausman test there is a maximum number of parameters which is the number of different equations that are used to obtain the two estimators. Three illustrative examples are used, namely an exogeneity test for the linear regression model, a test for the Box-Cox transformation, and a test for sample selection bias.

Keywords: Hausman test, specification test, number of parameters, instrumental variable (IV) model, Box-Cox model, Sample selection bias.

JEL classifications: C2; C5; I18

1. Introduction

Hausman (1978) developed a widely-used model specification test that has passed the test of time. The test is based on two estimators, one being consistent under the null hypothesis but inconsistent under the alternative, and the other being consistent under both the null and alternative hypotheses.

The difference of two estimators and the corresponding variance are used to calculate the test statistic, which asymptotically follows the chi-squared distribution with degrees of freedom given by the number of parameters. Holly (1982, p. 749) wrote “Hausman's procedure seems to be more general than the classical procedures for it does not seem to require that the null hypothesis be given in a parametric form.”

This paper considers the Hausman test for a case in which two estimators are obtained as roots of two different sets of equations. Some equations of the two sets may be the same, but at least one equation is different. The null hypothesis is that two estimators obtained from different sets of equations converge to the same values. It is shown that it may not be possible to use all the parameters in the model for the Hausman test. The asymptotic variance of the difference of the two estimators may converge to a singular matrix, and there exists a maximum number of parameters that can be used in the Hausman test. The maximum number of parameters that can be used in the test is determined by the number of different equations in the two sets. This result coincides with the case of a standard parametric test, where the degrees of freedom are given by the number of restrictions in the null hypothesis.

The remainder of the paper is given as follows. A Hausman test for a general model is discussed in Section 2, and three illustrative examples are given in Section 3.

2. A Hausman Test for a General Model

Let θ be a k -dimensional vector of unknown parameters. Consider two estimators, $\hat{\theta}$ and $\tilde{\theta}$, where $\hat{\theta}$ is consistent under the null hypothesis and inconsistent under the alternative, whereas $\tilde{\theta}$ is consistent under both the null and alternative hypotheses. Suppose that $\hat{\theta}$ is the root of k equations given by:

$$f_T(\theta) = 0, \text{ and } g_T(\theta) = 0 \quad (1)$$

where $f_T(\theta) = 0$ and $g_T(\theta) = 0$ express vectors of k_1 and k_2 different equations, respectively. On the other hand, $\tilde{\theta}$ is given by:

$$f_T(\theta) = 0, \text{ and } h_T(\theta) = 0. \quad (2)$$

Standard conditions such as convergence and differentiability of functions are assumed to hold. All models, whether linear or nonlinear, that are estimated using moment restrictions (by the method of moments) belong to this category.

Let θ_0 be the true parameter value of θ . Under the null hypothesis, it follows that:

$$\begin{aligned} f_T(\hat{\theta}) &= f_T(\theta_0) + f_T'(\theta_0)(\hat{\theta} - \theta_0) + o_p(1/\sqrt{T}), \\ f_T(\tilde{\theta}) &= f_T(\theta_0) + f_T'(\theta_0)(\tilde{\theta} - \theta_0) + o_p(1/\sqrt{T}). \end{aligned} \quad (3)$$

Since $f_T(\hat{\theta}) = 0$ and $f_T(\tilde{\theta}) = 0$, it follows that:

$$f_T'(\theta_0)\sqrt{T}(\hat{\theta} - \tilde{\theta}) = o_p(1). \quad (4)$$

This means that there are k_1 linear relations between $\sqrt{T}\hat{\theta}$ and $\sqrt{T}\tilde{\theta}$ asymptotically, and only K_2 elements are linearly independent asymptotically. Let $\hat{\theta}^*$ and $\tilde{\theta}^*$ be subsets of $\hat{\theta}$ and $\tilde{\theta}$ (in order to choose the corresponding elements), and let q be their dimension. If

$q > k_2$, $T \cdot V(\hat{\theta}^* - \tilde{\theta}^*)$ converges to a singular matrix under the null hypothesis, so that we cannot use the Hausman test in this case.

Note that another alternative is to use the generalized inverse matrix instead of $V(\hat{\theta}^* - \tilde{\theta}^*)^{-1}$. However, as the asymptotic distribution may not be the chi-squared distribution, this alternative is not recommended unless the asymptotic distribution is known.

3. Illustrative Examples

In this section we give three illustrative examples of the Hausman test, namely an exogeneity test for the linear regression model, a test for the Box-Cox transformation, and a test for sample selection bias.

3.1 An exogeneity test for the linear regression model

As the first example, we consider a classical exogeneity test of the linear regression model, as given by:

$$y_t = x_{1t}'\beta_1 + x_{2t}'\beta_2 + u_t = x_t'\beta + u_t, \quad t = 1, 2, \dots, T, \quad (5)$$

where $x_t' = (x_{1t}', x_{2t}')$, $\beta' = (\beta_1', \beta_2')$, x_{1t} is the k_1 'th dimensional vectors of the explanatory variables which is known to satisfy $\text{cov}(x_{1t}, u_t) = 0$, and x_{2t} is the k_2 'th dimensional vectors of the explanatory variables which might be $\text{cov}(x_{2t}, u_t) \neq 0$. We assume that all the other standard assumptions of the explanatory variables and error terms are satisfied.

For this model, we consider the test where the null and alternative hypotheses are given by:

$$H_0 : \text{cov}(x_{2t}, u_t) = 0, \quad H_1 : \text{cov}(x_{2t}, u_t) \neq 0. \quad (6)$$

This test is a classical example of the Hausman test, and has been examined extensively (see, for example, Durbin (1954), Wu (1973), Smith (1983, 1984, 1985), Holly (1982), and Hausman and Taylor (1980, 1981). However, the problem has not been examined in the context of this paper, where we can reach the conclusion much more simply than using existing methods. (As it is possible to treat the procedure as a standard parametric test, it may not be a good example of Holly's (1982) statement that "it does not seem to require that the null hypothesis be given in a parametric form").

It is well known that, under the null hypothesis, the ordinary least squares (OLS) estimator is consistent and efficient if the error terms are independently and identically distributed (i.i.d.) normal random variables. However, OLS is inconsistent under the alternative hypothesis. On the other hand, the IV estimator, such as the two-stage least squares (2SLS) estimator, is consistent under both the null and alternative hypotheses.

Let $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}$ be the OLS estimators of β_1 , β_2 and β , $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}$ be the IV estimators and $k = k_1 + k_2$. The OLS estimator is given by:

$$\sum_t x_{1t}(y_t - x_{1t}'\beta_1 - x_{2t}'\beta_2) = 0, \text{ and } \sum_t x_{1t}(y_t - x_{1t}'\beta_1 - x_{2t}'\beta_2) = 0, \quad (7)$$

and the IV estimator is given by:

$$\sum_t x_{1t}(y_t - x_{1t}'\beta_1 - x_{2t}'\beta_2) = 0, \text{ and } \sum_t z_{1t}(y_t - x_{1t}'\beta_1 - x_{2t}'\beta_2) = 0, \quad (8)$$

where z_t is a vector of variables which satisfies $\sum_t z_{1t}u/T \xrightarrow{P} 0$. The first k_1 equations are the same, and the differences arise in the latter k_2 equations. As the first k_1 equations yield the OLS estimators, we have:

$$\hat{\beta}_1 = \sum_t (x_{1t}x_{1t}')^{-1} \sum_t \{x_{1t}(y_t - x_{2t}'\hat{\beta}_2)\}, \text{ and } \tilde{\beta}_1 = \sum_t (x_{1t}x_{1t}')^{-1} \sum_t \{x_{1t}(y_t - x_{2t}'\tilde{\beta}_2)\} \quad (9)$$

and

$$\hat{\beta}_1 - \tilde{\beta}_1 = \left\{ \sum_t (x_{1t} x_{it}')^{-1} \sum_t (x_{1t} x_{2t}') \right\} (\tilde{\beta}_2 - \hat{\beta}_2), \quad (10)$$

which is a linear function of $\hat{\beta}_2 - \tilde{\beta}_2$. Therefore, if we choose $q > k_2$, $V(\hat{\beta}^* - \tilde{\beta}^*)$ becomes a singular matrix. If u_t and x_t are independent, the variance matrix become singular even in finite samples. As the structure of the model is simple, we can use the Moore-Penrose general inverse for the test (Hausman and Taylor, 1980) if $q > k_2$ in this particular case.

3.2 A test for the Box-Cox transformation

The second example is the Box-Cox (1964) transformation model (BC model), which is given by:

$$z_t = x_t' \beta + u_t, \quad y_t \geq 0, \quad t = 1, 2, \dots, T, \quad (11)$$

$$z_t = \frac{y_t^\lambda - 1}{\lambda}, \quad \text{if } \lambda \neq 0,$$

$$z_t = \log(y_t), \quad \text{if } \lambda = 0,$$

Generally, the likelihood function under the normality assumption (BC likelihood function) is misspecified, and the maximum likelihood estimator (BC MLE) is not consistent. However, the BC MLE can be a consistent estimator under certain assumptions. Nawata (2013) proposed an estimator which is consistent even if the assumption is not satisfied. Therefore, we can use the Hausman test for this model using these estimators.

We will explain the asymptotic distribution of Nawata's estimator and then the asymptotic distribution of the BC MLE. It is also shows that the Hausman test holds just for the transformation, and cannot be used for more than two parameters.

i) **BC MLE**

The likelihood function under the normality assumption (BC likelihood function) is given by:

$$\log L(\theta) = \sum_t \left[\log \phi\{(z_t - x_t' \beta) / \sigma\} - \log \sigma \right] + (\lambda - 1) \sum_t \log y_t, \quad (12)$$

where ϕ is the probability density function of the standard normal assumption, and σ^2 is the variance of u_t . The BC MLE is obtained as follows:

$$\begin{aligned} \frac{\partial \log L}{\partial \beta} &= \frac{1}{\sigma^2} \sum_t x_t (z_t - x_t' \beta) = 0, \quad \frac{\partial \log L}{\partial \sigma^2} = \sum_t \frac{(z_t - x_t' \beta)^2 - \sigma^2}{2\sigma^4} = 0, \text{ and} \quad (13) \\ \frac{\partial \log L}{\partial \lambda} &= 0. \end{aligned}$$

Generally, the likelihood function is misspecified, and the maximum likelihood estimator (BC MLE) is not consistent. However, if $\lambda_0 \sigma_0 / (1 + \lambda_0 x_t' \beta_0) \rightarrow 0$ and $P[y_t < 0] \rightarrow 0$ (in practice, $P[y_t < 0]$ is small enough), the BC MLE performs well. Following Bickel and Doksum (1981), we call this the “small σ ” assumption, such that:

$$\frac{\partial \log L}{\partial \lambda} \Big|_{\theta_0} = -\frac{1}{\sigma_0^2 \lambda_0} \sum_t \{y_t^{\lambda_0} \log(y_t) - z_t^* u_t\} + \sum_t \log(y_t), \quad (14)$$

where $z_t^* = y_t^{\lambda_0}$. Under the “small σ ” assumption, (that is, $|\lambda_0 u_t / (\lambda_0 x_t' \beta_0 + 1)|$ is small) and $\lambda_0 \neq 0$, we have:

$$\begin{aligned} \log(y_t) &= \frac{1}{\lambda_0} \log(\lambda_0 x_t' \beta_0 + 1 + \lambda_0 u_t) = \frac{1}{\lambda_0} \left\{ \log(\lambda_0 x_t' \beta_0 + 1) + \log\left(1 + \frac{\lambda_0 u_t}{\lambda_0 x_t' \beta_0 + 1}\right) \right\} \\ &\approx \frac{1}{\lambda_0} \log(\lambda_0 x_t' \beta_0 + 1) + \frac{u_t}{\lambda_0 x_t' \beta_0 + 1}. \end{aligned} \quad (15)$$

Therefore,

$$\begin{aligned} \frac{\partial \log L}{\partial \lambda} \Big|_{\theta_0} \approx & -\frac{1}{\sigma_0^2 \lambda_0} \sum_t \left[\left\{ \frac{(\lambda_0 x_t' \beta_0 + 1) \log(\lambda_0 x_t' \beta_0 + 1)}{\lambda_0} - x_t' \beta_0 \right\} u_t + \log(\lambda_0 x_t' \beta_0 + 1) u_t^2 + \frac{\lambda_0 u_t^3}{\lambda_0 x_t' \beta_0 + 1} \right] \\ & + \sum_t \left\{ \frac{1}{\lambda_0} \log(\lambda_0 x_t' \beta_0 + 1) + \frac{u_t}{\lambda_0 x_t' \beta_0 + 1} \right\}. \end{aligned} \quad (16)$$

Hence the BC MLE becomes a consistent estimator and the “small σ asymptotics” of the BC MLE $\hat{\theta}_{BC} = (\hat{\beta}_{BC}, \hat{\sigma}_{BC}^2, \hat{\lambda}_{BC})$ are obtained by:

$$\sqrt{T}(\hat{\theta}_{BC} - \theta_0) \rightarrow N(0, A^{-1}BA^{-1}), \quad (17)$$

where $A = -E\left[\frac{\partial^2 \log L}{\partial \theta \partial \theta'} \Big|_{\theta_0}\right]$, $B = E[\ell_t(\theta_0)\ell_t(\theta_0)']$, $\ell_t(\theta)' = [\xi_t(\theta)', \varsigma_t(\theta), g_t(\theta)]$,

$$\xi_t(\theta) = \frac{1}{\sigma^2} x_t(z_t - x_t' \beta), \text{ and } \varsigma_t(\theta) = \frac{(z_t - x_t' \beta) - \sigma^2}{2\sigma^2}.$$

ii) Nawata's estimator

Nawata (2013) proposed an estimator which is consistent even if the “small σ ” assumption is not satisfied. This estimator is obtained by:

$$\begin{aligned} \frac{\partial \log L}{\partial \beta} = 0, \quad \frac{\partial \log L}{\partial \sigma^2} = 0, \text{ and} \quad (18) \\ G_T(\theta) = \sum_t \left[-\frac{1}{\sigma^2 \lambda} \left\{ \frac{\log(\lambda x_t' \beta + 1)}{\lambda} + \frac{z_t - x_t' \beta}{\lambda x_t' \beta + 1} \right\} y_t^\lambda - z_t \right] (z_t - x_t' \beta) \\ + \frac{1}{\lambda} \log(\lambda x_t' \beta + 1) + \frac{z_t - x_t' \beta}{\lambda x_t' \beta + 1} \Big] \equiv \sum_t g_t(\theta) = 0, \end{aligned}$$

$G_T(\theta)$ is obtained by the approximation of $\partial \log L / \partial \lambda$. If the first and third moments of u_t are zero, we have $E[G_T(\theta_0)] = 0$ and the estimator obtained by equation (18) is consistent. The asymptotic distribution of the estimator, $\hat{\theta}_N' = (\hat{\beta}_N', \hat{\sigma}_N^2, \hat{\lambda}_N)$, is given by:

$$\sqrt{T}(\hat{\theta}_N - \theta_0) \rightarrow N[0, C^{-1}B(C')^{-1}], \quad (19)$$

where $C = -E[\frac{\partial \ell_t(\theta)}{\partial \theta'} |_{\theta_0}]$.

iii) The Hausman test for the “small σ ” assumption

In this case, the null hypothesis is that the “small σ ” assumption holds, and the alternative hypothesis is that it does not. As the variance of $V(\hat{\theta}_{BC} - \hat{\theta}_C)$ is given by $V(\hat{\theta}_{BC} - \hat{\theta}_C) = (A^{-1} - C^{-1})B\{A^{-1} - (C')^{-1}\}$ under the null hypothesis, we can calculate it more precisely than using the difference of two covariance matrices. The asymptotic distribution of $\hat{\lambda}_N - \hat{\lambda}_{BC}$ is given by:

$$\sqrt{T}(\hat{\lambda}_N - \hat{\lambda}_{BC}) \rightarrow N(0, d), \quad (20)$$

where

$$d = p \lim_{n \rightarrow \infty} T \cdot [\text{last diagonal element of } (A^{-1} - C^{-1})B^{-1}\{A^{-1} - (C')^{-1}\}].$$

Using $t = \sqrt{T}(\hat{\lambda}_N - \hat{\lambda}_{BC}) / \sqrt{\hat{d}}$ as the test statistic, where \hat{d} is the estimator of d , we can test the “small σ ” assumption, that is, we can test whether we can successfully use the BC MLE. (Note that, when $\lambda_0 = 0$, we replace $\lim_{\lambda_0 \rightarrow 0} A$, $\lim_{\lambda_0 \rightarrow 0} B$, and $\lim_{\lambda_0 \rightarrow 0} C$ for A , B and C , and the test can be calculated using the same formula.)

However, in this case:

$$\hat{\beta}_{MLE} = \left\{ \sum_t (x_t x_t')^{-1} \right\} \left\{ \sum_t x_t y_t^{\hat{\lambda}_{MLE}} \right\}, \text{ and } \hat{\beta}_N = \left\{ \sum_t (x_t x_t')^{-1} \right\} \left\{ \sum_t x_t y_t^{\hat{\lambda}_N} \right\} \quad (21)$$

$$\begin{aligned} \text{as } \hat{\beta}_{MLE} - \hat{\beta}_N &= \left\{ \sum_t (x_t x_t')^{-1} \right\} \left\{ \sum_t x_t (y_t^{\hat{\lambda}_{MLE}} - y_t^{\hat{\lambda}_N}) \right\}, \\ &\left\{ \sum_t (x_t x_t')^{-1} \right\} \left\{ \sum_t x_t (\log y_t) y_t^{\lambda_0} \right\} (\hat{\lambda}_{MLE} - \hat{\lambda}_N) + o_p(1/\sqrt{n}), \\ \sqrt{T}(\hat{\beta}_{MLE} - \hat{\beta}_N) &= A^{-1}B \cdot \sqrt{T}(\hat{\lambda}_{MLE} - \hat{\lambda}_N) + o_p(1), \end{aligned} \quad (25)$$

where

$$A = p \lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_t x_t' \text{ and } B = p \lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_t y_t^{\lambda_0}.$$

Therefore, the rank of the asymptotic variance matrix of $[\sqrt{T}(\hat{\lambda}_{MLE} - \hat{\lambda}_N), \sqrt{T}(\hat{\beta}_{MLE} - \hat{\beta}_N)']$ becomes one, and we cannot add any element of β to the test.

3.3 A test for sample selection bias

In a sample selection bias model (Heckman, 1976, 1979), the equation to be estimated is given as:

$$y_t = x_{1t}' \beta_1 + \gamma G(x_{2t}' \beta_2) + u_t, \quad t = 1, 2, \dots, T, \quad (26)$$

The null hypothesis is that the error term follows the normal distribution, and $G(x_{2t}' \beta_2) = G_1(x_{2t}' \beta_2) \equiv \phi(x_{2t}' \beta_2) / \Phi(x_{2t}' \beta_2)$, where Φ is the distribution function of the standard normal distribution. We can estimate the model by Heckman's two-step method and obtain an estimator that satisfies:

$$\sum_t x_{1t}' \{y_t - x_{1t}' \hat{\beta}_1 - \hat{\gamma} G_1(x_{2t}' \hat{\beta}_2)\} = 0. \quad (27)$$

If the distribution is not known, we can estimate $G(x_t'\beta_2)$ by using a semiparametric method (such as, for example, the multiple index estimator of Ichimura and Lee (1991)). In this case, the estimator of β_1 is obtained from:

$$\sum_t x_{1t} \{y_t - x_{1t}'\tilde{\beta}_1 - \tilde{\gamma}G_{2T}(x_{2t}'\tilde{\beta}_2)\} = 0. \quad (28)$$

Moreover, under the null hypothesis, $G_{2T}(\xi)$ must satisfy $g_T(\xi) = G_{2T}(\xi) - G_1(\xi) = o_p(1/\sqrt{T})$ in a proper subspace of $R = (-\infty, \infty)$ to obtain an estimator such that the distribution of $\sqrt{T}(\tilde{\beta}_1 - \beta_1)$ converges to a normal distribution with mean zero. Therefore, under the null hypothesis, it follows that:

$$\begin{aligned} (\sum_t x_{1t}x_{1t}')(\tilde{\beta}_1 - \hat{\beta}_1) &= \sum_t \{\hat{\gamma}G_1(x_{2t}'\hat{\beta}_2) - \tilde{\gamma}G_{2T}(x_{2t}'\tilde{\beta}_2)\}x_{1t} \\ &= \sum_t \{\hat{\gamma}G_1(x_{2t}'\hat{\beta}_2) - \tilde{\gamma}G_1(x_{2t}'\tilde{\beta}_2) + \tilde{\gamma}G_1(x_{2t}'\tilde{\beta}_2) - \tilde{\gamma}G_1(x_{2t}'\tilde{\beta}_2)\}x_{1t} \\ &= \sum_t \{\hat{\gamma}G_1(x_{2t}'\hat{\beta}_2) - \tilde{\gamma}G_1(x_{2t}'\tilde{\beta}_2) + \tilde{\gamma}G_1(x_{2t}'\hat{\beta}_2) - \tilde{\gamma}G_1(x_{2t}'\tilde{\beta}_2) + \tilde{\gamma}g_T(x_{2t}'\tilde{\beta}_2)\}x_{1t} \\ &= \sum_t [G_1(x_{2t}'\hat{\beta}_2)(\hat{\gamma} - \tilde{\gamma}) + \tilde{\gamma}\{G_1(x_{2t}'\hat{\beta}_2) - G_1(x_{2t}'\tilde{\beta}_2)\} + \tilde{\gamma}g_T(x_{2t}'\tilde{\beta}_2)]x_{1t} \\ &= \sum_t \{G_1(x_{2t}'\beta_2)(\hat{\gamma} - \tilde{\gamma}) + \gamma \cdot g_1(x_{2t}'\beta_2)x_{2t}'(\hat{\beta}_2 - \tilde{\beta}_2) + \gamma \cdot g_T(x_{2t}'\beta_2)\}x_{1t} + O_p(1/\sqrt{T}), \end{aligned} \quad (29)$$

where $g_T(\xi) = G_1(\xi) - G_T(\xi)$, and $g_1(\xi) = \frac{dG_1}{d\xi}$, so that:

$$\sqrt{T}(\tilde{\beta}_1 - \hat{\beta}_1) = A^{-1}B\sqrt{T}(\hat{\gamma} - \tilde{\gamma}) + A^{-1}C\sqrt{T}(\hat{\beta}_2 - \tilde{\beta}_2) + o_p(1), \quad (30)$$

$$A = p \lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_{1t}x_{1t}', B = p \lim_{T \rightarrow \infty} \frac{1}{T} \sum_t G(x_{2t}'\beta_2)x_{1t}, \text{ and } C = p \lim_{T \rightarrow \infty} \frac{1}{T} \sum_t g_1(x_{2t}'\beta_2)x_{1t}.$$

Again, we have k_1 asymptotic restrictions for $(\beta_1', \beta_2', \gamma)$, which leads to the identical problem that was mentioned above. Note that Heckman's estimator under the normality assumption is consistent but not efficient. Therefore, we cannot use Hausman's simple

formula, whereby the variance in the test statistic is the difference of two covariance matrices, to calculate the variance of the test statistic.

4. Conclusion

In this paper, the Hausman (1978) test was re-examined. The widely-used model specification test uses two estimators, one being consistent under the null hypothesis but inconsistent under the alternative, and the other being consistent under both the null and alternative hypotheses. It was shown that the asymptotic variance of the difference of the two estimators could be a singular matrix. Moreover, in calculating the Hausman test there is a maximum number of parameters which is the number of different equations that are used to obtain the two estimators. Three illustrative examples are used, namely an exogeneity test for the linear regression model, a test for the Box-Cox transformation, and a test for sample selection bias.

The limitation of the Hausman test that was established in the paper does not seem to have been considered previously, except in the case of the classical exogeneity test for the linear regression model. This result suggests that greater care and attention should be paid in computing the Hausman test for such problems.

References

- Bickel, P.J. and K.A. Doksum (1981) "An Analysis of Transformations Revisited," *Journal of the American Statistical Association* **76**, 296-311.
- Box, G.E.P., and D.R. Cox (1964) "An Analysis of Transformations," *Journal of the Royal Statistical Society B* **26**, 211-252.
- Durbin, J. (1954) "Errors in Variables," *Review of the International Statistical Institute*, **22**, 23-32.
- Hausman, J. (1978) "Specification Tests in Econometrics", *Econometrica*, **46**, 1251-72.
- Hausman, J.A. and W. Taylor (1980) "Comparing Specification Tests and Classical Tests," MIT Department of Economics Discussion Paper No. 226.
- Hausman, J.A. and W. Taylor (1981) "A Generalized Specification Test," *Economics Letters*, **8**, 239-245.
- Heckman, J. (1976) "The Common Structure of Statistical Models of Truncation, Sample Selection Bias and Limited Dependent Variables and a Simple Estimator for such Models," *Annals of Econometric and Social Measurement*, **5**, 475-492.
- Heckman, J. (1979) "Sample Selection Bias as a Specification Error," *Econometrica*, **47**, 153-161.
- Holly, A. (1982) "A Remark on Hausman's Specification Test," *Econometrica*, **50**, 749-759.
- Ichimura, H. and L.-F. Lee (1991), "Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation", in *International Symposia in Economic Theory and Econometrics*, W.A. Barnett, J. Powell, and G. Tauchen (eds.), Cambridge University Press, pp. 3-49.
- Nawata K. (2013). "A New Estimator of the Box-Cox Transformation Model Using Moment Conditions," *Economics Bulletin*, **33**, 2287-2297.
- Smith, R. J., (1983) "On the Classical Nature of the Wu-Hausman Statistics for the Independence of Stochastic Regressors and Disturbance," *Economics Letters*, **11**, 357-364.
- Smith, R. J., (1984) "A Note on Likelihood Ratio Tests for the Independence Between a Subset of Stochastic Regressors and Disturbances," *International Economic Review*, **25**, 263-269.
- Smith, R.J. (1985) "Wald Tests for the Independence of Stochastic Variables and Disturbance of a Single Linear Stochastic Simultaneous Equation," *Economics Letters*, **17**, 87-90.
- Wu. D.M. (1973) "Alternative Tests of Independence Between Stochastic Regressors and Disturbances," *Econometrica*, **41**, 733-750.